

Van Landeghem, Bert

Working Paper

Panel conditioning and self-reported satisfaction: Evidence from international panel data and repeated cross-sections

SOEPPapers on Multidisciplinary Panel Data Research, No. 484

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Van Landeghem, Bert (2012) : Panel conditioning and self-reported satisfaction: Evidence from international panel data and repeated cross-sections, SOEPPapers on Multidisciplinary Panel Data Research, No. 484, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/65655>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SOEPpapers

on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel Study at DIW Berlin

484-2012

Panel Conditioning and Self-reported Satisfaction: Evidence from International Panel Data and Repeated Cross-sections

Bert Van Landeghem

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
<http://www.diw.de/soeppapers>

Editors:

Jürgen **Schupp** (Sociology, Vice Dean DIW Graduate Center)
Gert G. **Wagner** (Social Sciences)

Conchita **D'Ambrosio** (Public Economics)
Denis **Gerstorff** (Psychology, DIW Research Professor)
Elke **Holst** (Gender Studies)
Frauke **Kreuter** (Survey Methodology, DIW Research Professor)
Martin **Kroh** (Political Science and Survey Methodology)
Frieder R. **Lang** (Psychology, DIW Research Professor)
Henning **Lohmann** (Sociology, DIW Research Professor)
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Professor)
Thomas **Siedler** (Empirical Economics)
C. Katharina **Spieß** (Empirical Economics and Educational Science)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

Panel Conditioning and Self-reported Satisfaction: Evidence from International Panel Data and Repeated Cross-sections

Bert Van Landeghem*

Maastricht University, ROA

`b.vanlandeghem@maastrichtuniversity.nl`

August 2012

Abstract

Using data from three European countries, this paper investigates whether self-reported satisfaction data are subject to panel conditioning or a panel effect, that is, whether answers depend on whether one has previously participated in the panel. The analysis proposes a way to account for panel attrition in cases where the attrition rate is substantial, and finds international evidence for a negative panel effect. This result can be of importance e.g. in the current debate on trends in life satisfaction over time), and might stimulate further research into panel conditioning for other economic data gathered via household surveys.

keywords: Life Satisfaction, GHQ, Panel Conditioning, Panel Attrition, Time Trends

JEL Codes: C83, I30, J10

*I thank Eric Bonsang, Arnaud Dupuy, Richard Easterlin, Arjan Non, Andrew Oswald and Arthur van Soest for useful comments and discussions. I am also grateful to seminar and conference participants in Berlin, Bonn, Cambridge, Freiburg, Maastricht and Milan.

1 Introduction

During the past several decades, economists have paid greater attention to ‘stated utility’, as opposed to ‘revealed preferences’. While the latter are inferred from economic agents’ actions, the former are derived from simple survey questions, stated as follows:

On a scale from 0 (very unsatisfied) to 10 (very satisfied), how satisfied are you with your life, all things considered?

Increasing evidence underpins the internal validity of such data (Krueger and Schkade, 2008; Sgroi et al., 2010), as well as the external validity (Oswald and Wu, 2010). Researchers have examined many relationships between happiness¹ data and socioeconomic variables, and whether these should be interpreted as causal. Happiness has been shown to be strongly correlated with relational goods and social capital (Becchetti et al., 2008; Powdthavee, 2008), with major life events such as unemployment, bereavement, and disability (Clark et al., 2007; Oswald and Powdthavee, 2008a,b), with social status (Blanchflower et al., 2009; Luttmer, 2005; Ravallion and Lokshin, 2010; Senik, 2004, 2008a, 2009), and with expectations and aspirations (de Grip et al., 2012; McBride, 2010; Senik, 2008b). Others studies have analyzed geographical differences in life satisfaction (Oswald and Wu, 2011) or to what extent happiness is determined by genetics (De Neve et al., 2010). Very recently, economists are changing their focus from the causes of happiness to what happiness itself can cause. De Neve and Oswald (2012), Oswald et al. (2012) and Proto et al. (2012), find a causal impact of happiness on productivity, while Ifcher and Zarghamee (2011) find a negative causal link between happiness and time preference.

In fact, subjective well-being data have become so crucial in economic research that even in mainstream economic journals, methodological issues regarding such data are repeatedly being discussed. Bertrand and Mullainathan (2001) and Conti and Pudney (2011) argue that subtle changes in question wording and questionnaire design can influence results, not only the average scores in the raw data but also the correlations with socioeconomic variables. Oswald (2008) argues that one’s self-reported life satisfaction might be a concave function of *true* life satisfaction. This would imply that the log-linear relationship, which is often found between income and reported well-being, does not necessarily reflect the functional relationship between income and true well-being.

¹Terms such as happiness, life satisfaction, and subjective well-being are often used interchangeably in the literature.

Ferrer-i-Carbonell and Frijters (2004) find that treating happiness data as cardinal rather than as ordinal has no substantial impact on regression results, but accounting for fixed effects has a dramatic impact on the results. Others have been concerned about anchoring effects, that is, people do not report an absolute level of happiness but rather their happiness relative to a reference point. Recently, these concerns have been addressed using vignettes, describing hypothetical individuals or households. The ranking of these vignettes on a happiness ladder by respondents can then give us some information about heterogeneity in anchors across individuals or countries. Such investigations have led to mixed results. Beegle *et al.* (2012), using data from the Tajikistan household survey conclude that, although it seems that individuals have different benchmarks, these do not seem to significantly affect the estimated relationship with objective economic indicators. Bonsang and van Soest (2012) use vignettes to address differences in anchoring across countries and find for older individuals in 11 European countries that cross-country differences in subjective well-being are more in line with cross-country differences in income, after correcting scores with vignettes. In a study comparing the Netherlands and the United States, Kapteyn *et al.* (2011) use vignettes to correct for anchoring effects both within and across countries. They find that correcting subjective well-being scores with vignettes leads to almost identical distributions of subjective well-being in both countries and, additionally, leads to a substantially higher correlation between subjective well-being and income, particularly in the United States.

Another methodological issue about which one might be concerned is a phenomenon called ‘panel conditioning’ or a ‘panel effect’, which implies that answers to questions depend on whether one has previously participated in the panel. Das *et al.* (2011) and Toepoel *et al.* (2009) offer a cross-disciplinary overview of the literature on panel conditioning, and by comparing two samples, one fresh and the other, more experienced, they find that a panel effect is especially strong for knowledge questions but not for attitudinal questions. The authors note, however, that research on panel conditioning is still rather limited and that generally no attempt is made to distinguish panel conditioning from panel attrition. In the context of life satisfaction measures, both Frijters and Beaton (2012), and Kassenboehmer and Haisken-Denew (2012), who try to explain the frequently found U-shaped pattern of subjective well-being over the life cycle, find a negative coefficient on a variable measuring the number of times having participated into the panel. However, apart from potential attrition biases, in a model that includes individual fixed effects, age, and time effects, it might be difficult to interpret a variable measuring the length of staying in the panel. Indeed, multicollinearity problems force

researchers (often unconsciously) to make arbitrary identifying assumptions; these can make a dramatic impact on results and hence explain why recent life-course studies using the same data sometimes report different outcomes (see e.g. Clark and Oswald, 2006 vs. Frijters and Beatton, 2012, and Gerstorf et al., 2010). In psychology, in contrast, Sharpe and Gilbert (1998) find that testing individuals for depression twice with a one-week interval leads to a decrease in self-reported negative emotions and does not seem to have an effect on self-reported positive emotions. However, the very specific small sample of college undergraduates, as well as the very short time period that has collapsed between the two sessions, make it difficult to extrapolate the conclusions to nation-wide panel surveys repeated with yearly intervals.

This paper investigates the existence of panel conditioning in several ways. Reported life satisfaction of refreshment samples is compared with reported life satisfaction of more experienced samples within the same panel dataset. Issues of panel ageing and panel attrition, which might cloud results, will be addressed. As a robustness check, trends in panel data are compared with trends in repeated cross-sections within the same country.

The findings of this paper might be useful to the growing literature studying trends of subjective well-being over time. Blanchflower and Oswald (2004) find happiness runs flat over time in Britain, and in the United States, it is increasing for blacks but decreasing for white women. Stevenson and Wolfers (2009) document how in the United States, over the last three decades, women's happiness has decreased both in absolute terms as well as relative to men and that nowadays men are even happier than women, whereas it was the reverse three decades ago. They find an upward-sloping trend in life satisfaction for both European men and women, although the increase was smaller for women than for men. Indeed, if self-reported life satisfaction data are subject to panel conditioning, this would imply that trends derived from panel data are likely to be clouded and, at the very least, panel data and repeated cross-sections cannot be used interchangeably. To derive trends, Blanchflower and Oswald (2004), for example, use repeated cross-sections for the United States, while for Britain they use panel data. Stevenson and Wolfers (2009) use repeated cross-sections during their analysis but refer to an unreported robustness check in German panel data.

A panel effect's existence could also be important for researchers interested in how macroeconomic indicators are correlated with subjective well-being (Di Tella et al. (2001, 2003) and Wolfers (2003)). In particular, one can think of the renewed debate on the bivariate (absence of) correlation between economic growth and happiness, a topic dis-

cussed in works such as Clark and Senik (2010), Deaton (2008), Di Tella and MacCulloch (2008), Easterlin (1974, 1995, 2001), Easterlin et al. (2010), Sacks et al. (2010), and Stevenson and Wolfers (2008).

The remainder of the paper is organized as follows. Section Two documents the data that will be used and outlines the core of the empirical strategy. Section Three presents and discusses the econometric results, and Section Four concludes.

2 Data and Empirical Strategy

2.1 Panel Data

The study draws upon three panel datasets, the German Socioeconomic Panel (SOEP), the British Household Panel Survey (BHPS), and the Swiss Household Panel (SHP). In all three panels, by definition, the aim is to re-interview the same individuals in successive rounds. New individuals can enter, however, for several reasons. For example, new samples can be added to the survey to oversample some minorities or to refresh the panel. Apart from adding new samples, each year, the datasets will contain a small number of new respondents. First, members from interviewed households will reach the eligible age to enter the panel. Second, if a new member eligible for the survey moves into the household, the enumerator will attempt to interview this new member as well. Third, if a household member leaves the household, the aim is to follow the respondent and, at the same time, try to interview other members eligible for the survey in that person's new household. Generally, new respondents not stemming from a refreshment sample account for around 3% of the respondents; they ensure that each age group is represented in each survey round.²

Among the three panel datasets analysed, the German SOEP is probably the one that allows the most extensive analysis and that is the most common panel dataset in happiness research. The German SOEP is provided by DIW Berlin and is repeated at yearly intervals, running from 1984 for West Germany and 1990 for East Germany (see Wagner et al., 2007).

²To enhance readability, the refreshers stemming from these three sources will be referred to as 'natural refreshers', while refreshers stemming from the introduction of a refreshment sample will be referred to as 'induced refreshers'.

The observations can be categorized into eight³ different samples:

- Sample A (started in 1984) represents the West German population, while sample C (started in 1990) represents the East German population.
- Samples E, F, and H (started in 1998, 2000 and 2006 respectively) are refreshment samples.
- Samples B and D (started in 1984 and 1994 respectively) are immigrant subsamples, and sample G (started in 2002) comprises high-income households. The latter three samples are excluded from the analysis to avoid that well-being trends are artificially influenced by the oversampling of one of these categories from a certain calendar year onwards.

This results in a sample of around 31,000 individuals: 21,000 in West Germany and 10,000 in East Germany, or 190,500 and 92,500 person-year observations, respectively.

The following question on subjective well-being, asked at the end of a face-to-face questionnaire, has been in the survey from 1984 onwards and ever since has reappeared in every survey round.

On a scale from 0 (completely dissatisfied) to 10 (completely satisfied): How satisfied are you with your life, all things considered?

Across all rounds, the average life satisfaction score in West Germany equals 7.15 with a standard error of 1.82. In East Germany, life satisfaction is, on average, considerably lower than in West Germany, with an average score of 6.41 and standard error of 1.80.

The BHPS is made available through the ESRC Data Archive and originally was collected by the ESRC Research Centre on Microsocial Change at the University of Essex. The BHPS was begun in 1991. In 1999, two additional samples were included to oversample households in Wales and Scotland. In 2001, a Northern Ireland sample was started, so the survey could cover the United Kingdom area rather than just Great Britain. From 1997 to 2001, additional households were interviewed for the purpose of the European Community Household Panel (ECHP), and those households that agreed their information could be passed on to the University of Essex constitute another subsample in the BHPS, the ECHP subsample (See Taylor et al., 2010 for a detailed description of the BHPS data). The four latter samples are, however, not included in the analysis, again

³In 2009, a ninth sample was started, but the analysis will not make use of this sample as these individuals are currently observed in only two rounds of the panel.

to avoid artificial results, leaving us with data for around 15,300 individuals and 77,000 person-year observations.

Questions about people's well-being are asked at the beginning of a self-completed questionnaire. Unfortunately, however, not all rounds contain all the information that should ideally be used throughout the analysis. A general life satisfaction question is only included from the year 1996 onwards and was skipped in 2001, which means that we have subjective well-being data only for the periods 1996-2000 and 2002-2008. The satisfaction-with-life data for waves 6-10 and 12-18 are based on the following survey question:

Please tick the number that you feel best describes how dissatisfied or satisfied you are with your life overall. 1 (not satisfied at all) to 7 (completely satisfied)

Across these 12 waves and all surveyed individuals, respondents report an average life satisfaction score of 5.22 on a scale of 7, with a standard deviation of 1.29. In 1.48% of the responses, the individual indicates being completely dissatisfied, while in 14.49% of cases, respondents are completely satisfied.

All 18 rounds, however, include questions from a General Health Questionnaire (GHQ), which are designed to identify depression. Due to a lack of subjective well-being data, the GHQ data are often used instead (e.g. Metcalfe et al., 2011). In particular, the following are queried:

Here are some questions regarding the way you have been feeling over the last few weeks. For each question, please tick the box next to the answer that best describes the way you have felt.

Have you recently...

1. been able to concentrate on whatever you're doing? ++
2. lost much sleep over worry? -
3. felt that you were playing a useful part in things? ++
4. felt capable of making decisions about things? ++
5. felt constantly under strain? -
6. felt you couldn't overcome your difficulties? -
7. been able to enjoy your normal day-to-day activities? ++

8. been able to face up to problems? ++
9. been feeling unhappy or depressed? –
10. been losing confidence in yourself? –
11. been thinking of yourself as a worthless person? –
12. been feeling reasonably happy, all things considered? ++

For the items followed by ‘++’ the response scale is as follows:

1. More so than usual
2. About the same as usual
3. Less so than usual
4. Much less than usual

And for the items followed by ‘–’:

1. Not at all
2. Not more than usual
3. Rather more than usual
4. Much more than usual

To construct an aggregate measure, a dummy is created for each item taking 1 when the score on the item equals 3 or 4. By summing the dummies within a person-year observation, one obtains a 0-12 depression scale. Subsequently, the score is reversed, so that 12 = no depressive symptoms and 0 = severely depressed. This measure (labelled as GHQ score) will be used as an alternative dependent variable during the regression analysis instead of the life satisfaction measure. Across all rounds and individuals, the average GHQ score equals 9.30, with a standard deviation of 3.20. Of respondents, 2.26% have the lowest GHQ score of 0, while 31.85% of respondents have the highest GHQ score of 12. Hence, there are considerably more individuals with the highest GHQ score than with the highest life satisfaction score, illustrating the upper truncation of the former measure. The GHQ score is appropriate to use in identifying different degrees of depression, but it cannot help us distinguish satisfied individuals from those who are extremely satisfied.

The SHP is a panel repeated at yearly intervals that started in 1999 and is run by the FORS, the Swiss Centre of Expertise in the Social Sciences. Currently, there

are 12 rounds available (up to 2010), with more than 95,000 person-year observations and 18,300 individuals. In 2004, a refreshment sample was started. Respondents are interviewed by telephone (See Voorpostel et al., 2011 for more details). A general life satisfaction question is asked from the year 2000 onwards, which proceeds as follows:

In general, how satisfied are you with your life if 0 means ‘not at all satisfied’ and 10 means ‘completely satisfied’?

The average score across all person-year observations equals 8.01 with a standard error of 1.48. Of those who answered, 16.5% consider themselves completely satisfied.

2.2 Repeated Cross-Sections

In order to compare results from panel data with results from repeated cross-sections, the paper draws upon data from the World Values Survey (WVS) and the Eurobarometer Survey.

The WVS is a face-to-face survey conducted worldwide that aims to facilitate investigating sociocultural and political change across regions and over time. The WVS now counts five waves spanning the periods 1981-1984, 1989-1993, 1994-1999, 1999-2004, and 2005-2008, respectively (World Values Survey, 2009). Currently, a sixth wave is being carried out. The WVS does not follow the same people over time. Nevertheless, many countries have been included in the survey in multiple rounds, offering us repeated cross-sectional data for them. In particular, the three countries studied in this paper have been surveyed at least twice, which means a comparison between panel data and repeated cross-sectional data is possible. There are data for Germany for the years 1997 and 2006, for Britain for the years 1998 and 2006, and for Switzerland for the years 1989, 1996, and 2007. For each country, around 1,100 to 2,000 individuals are surveyed in one round. The life satisfaction variable in the WVS is measured with the following question:

Taking all things together, would you say you are:

- Very happy
- Quite happy
- Not very happy
- Not at all happy

During the analysis, scores 1 and 4, and 2 and 3 are swapped so that 1 denotes ‘not at all happy’ and 4 denotes ‘very happy’, which makes the results easier to compare with results from the panel data. For the different waves and the three countries together, respondents report an average life satisfaction score of 3.11 (standard deviation 0.67). The highest happiness score is reported by 31.2% of respondents. This is a significantly larger percentage than in the panel datasets, but this might result from the more restricted scale that allows respondents less differentiation. Only 1.6% of individuals report the lowest happiness score.

A second repeated cross-sectional dataset is the Mannheim Eurobarometer trend file, which runs from 1970 to 1999. A variable is defined as a trend variable if it can be found in a Eurobarometer Survey (conducted with face-to-face interviews) in at least five survey years (Scholz and Schmitt, 2001). Compared with the WVS, this dataset has the advantage of containing cross-sections repeated at yearly intervals. A first subjective well-being question is:

On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?

As with the subjective well-being measure in the WVS, scores 1 and 4 are swapped, as well as scores 2 and 3, to obtain a subjective well-being variable going from low to high well-being. The question has been asked on a yearly basis from 1973 to 1998, except for the years 1974 and 1996.⁴ Averaging all available data for West and East Germany and Britain gives us a life satisfaction score of 3.13 with standard deviation 0.68. Of individuals, 2% report the lowest well-being score, while 29% report the highest well-being score.

2.3 Empirical Strategy

As is clear from Section 2.1, two out of the three panel datasets contain refreshment samples in the strict sense, that is, new samples drawn from the same population from which the original sample was drawn. Refreshment samples are often added in the course of a panel study to increase the sample size, allowing for a wider possibility of analyses, or at least to mitigate any decrease in sample size due to panel attrition.

⁴The Eurobarometer trend file also contains a happiness question. This question has no longer been asked after 1986, which means that these data cannot be used in comparison with the panel datasets.

Such refreshment samples can help us identify whether answers on (subjective well-being) survey questions are prone to panel conditioning. Indeed, they allow us to compare the average responses between a group of first-time respondents on the one hand and more experienced respondents on the other hand. When comparing a more experienced sample with refreshers within the same panel study, one need not worry about differences in survey design. The following cross-sectional regression specification can be estimated for each calendar year in which a refreshment sample was introduced:

$$W_{it} = \alpha_t + \beta_t R_{it} + e_{it} \quad (1)$$

Where the subscript i indexes the individual and t the calendar year, W denotes subjective well-being and R is an indicator taking the value of one when an individual i either belongs to a refreshment sample started in year t or happens to be a natural refresher belonging to the more experienced sample in that year. e_{it} is an individual-specific error term, and α_t and β_t parameters to be estimated. The coefficient β_t can be interpreted as the difference in well-being between the fresh and more experienced respondents in year t .

One may of course be sceptical as to whether the two groups of new and experienced respondents are comparable and if the results are clouded by panel attrition. One might consider the strategy of Das et al. (2011), who extend the framework of Keisuke et al. (2001) by recognizing that differences between an experienced sample and a refreshment sample can be due to not only panel attrition but also panel conditioning. Das et al. (2011) point out if no assumptions regarding the attrition process are to be made, one can still identify upper and lower bounds of the panel effect. They illustrate this with dichotomous responses and calculate bounds by assuming that all attritors would have chosen either 0 or 1. In this context, however, calculating bounds seems less useful. The range of possible ordinal responses is much larger than just two options, and moreover, the attrition rate in the considered datasets is relatively high (see below). However, since refreshment samples are all introduced well before the last calendar year for which data are available, one can rerun regressions as described in Equation (1), but now restricting the refreshment sample to those who will respond to the subjective well-being question in at least N consecutive survey rounds.

Subsequently, one might be interested in the actual size of the panel effect as well as in the dynamics. Is a panel effect of importance only between the first and second interview,

or does it cumulate across several interview rounds? After defining \bar{N}_i to be the total number of consecutive rounds in which the individual i has answered to the subjective well-being question, the following pooled cross-sectional regression specification could help us to gain more insight in these issues:

$$W_{it} | \bar{N}_i \geq N = \sum_{n=1}^N (\beta_n I_{nt}) + \sum_{t=2}^T (\gamma_t D_t) + \alpha + e_{it} \quad (2)$$

Again, W denotes subjective well-being. The subscripts i , t , and n are to index the individual, the calendar year and the number of subsequent interviews an individual has participated into the panel. $I_1, I_2 \dots I_N$ are indicators for the first, second, \dots , N th subsequent interview. The baseline category consists of those who are being interviewed more than N times. Some calendar years will have many more new respondents than average years, in particular those years in which a new sample was started. Hence, time dummies (D_2 to D_T) are included as controls to address the problem of nonrandom distribution of newcomers across calendar years. Since the regressions are run on a subsample of individuals who are in the sample for at least N consecutive survey rounds, we need not worry that the path of the panel effect over the different survey rounds (from the first to the N th) is clouded by panel attrition. Panel ageing might be an issue, since an individual answering the survey for the n th time can never be younger than the minimum age of a respondent plus $n - 1$. This should be a minor concern, though, when N is taken relatively small. Indeed, even though the literature does not entirely agree on how life satisfaction evolves through life, it seems that most researchers do agree that life cycle happiness evolves rather smoothly and not with big jumps (Blanchflower and Oswald, 2008; Easterlin, 2006; Fischer, 2009; Gwozdz and Sousa-Poza, 2010).

Finally, as a robustness check, the paper compares trends in panel data with trends in cross-sectional data within the same country. The regressions estimating trends in data are kept very parsimonious and contain no variables other than calendar year dummies. In particular, including age dummies and fixed effects, which are standard controls in happiness equations, would lead to estimates that are very hard to interpret, and seemingly mild parametric restrictions to obtain identification can lead to rotations of the estimated path of well-being over time if the underlying assumptions are not entirely correct (see e.g. Van Landeghem, 2012, Appendix 1 for an algebraic example).

3 Results

3.1 A Comparison between New Respondents and Experienced Respondents within a Panel

This subsection starts by documenting the difference between average subjective well-being scores of a refreshment sample and the average scores of more experienced respondents for that same year within the same panel dataset. In the years that the refreshment samples are started, each has a substantial number of new respondents, which facilitates the test. The German SOEP contains three refreshment samples in the years 1998, 2000, and 2006, while the SHP introduced a refreshment sample in 2004. For West Germany, the three refreshment samples (plus the natural refreshers who entered in the same years) offer us 1,747, 8,891, and 2,171 new respondents, respectively, while the figures are 449, 2,245, and 591 for East Germany. In the SHP, the 2004 panel refreshment (together with the natural refreshers in that year) bring us data on new respondents for 5,371 individuals.

The differences in average scores for the new respondents in the refreshment samples on the one hand and the experienced sample in the corresponding calendar year on the other hand are obtained by running regression specifications of the form as in Equation (1), and are presented in Table 1 for West Germany, East Germany, and Switzerland. Weights are used in order to ensure each age group has the same weight in both the refreshment and experienced sample.

In all cases, scores in the calendar year in which a refreshment sample is started are higher for the refreshment sample than for the more experienced sample, even after using weights to correct for differences in age distributions. The results are statistically significant at any significance level, and the magnitude is substantial, varying between 0.16 and 0.64 on a 0-10 scale. For completeness, Table 2 and 3 report results separately for men and women, respectively and show the results hold for both sexes. There might be concern the refreshment samples are nevertheless substantially different from the more experienced samples. For example, Table 4 shows us probit equations for each survey year in which a refreshment sample was started, and these predict whether or not an individual is a new respondent. The results from the regressions make clear that socioeconomic characteristics are not always completely randomly distributed across the refreshment and the experienced samples due to sampling error or attrition biases or

maybe because the explanatory variables are subject to panel conditioning themselves. The signs and size of the coefficients on the variables predicting the likelihood of belonging to a refreshment sample are, however, not very consistent across the different regressions. One common trend across the different regressions seems to be that singles are more likely to be in the refreshment sample than in the more experienced sample. This is not completely illogical, since singles are more mobile and thus more prone to quit the panel. It also seems that, across the different regressions, individuals with lower household income are more likely to be first-time respondents. We have a closer look at the distribution of this latter variable in Table 5, which shows the difference between real household income in the refreshment subsample on the one hand and in the experienced sample on the other hand, using weights to correct for differences in age distributions between both samples. The results suggest that differences in average income between the refreshment samples and experienced samples are rather small (see Table 5). Real household income in the refreshment samples for Germany are, on average, 0 to 17% lower than in the more experienced sample. As for the SHP, real household income is around 6% lower in the refreshment sample than in the experienced sample. These results thus seem to offer reassurance that the differences between well-being in refreshment and experienced samples shown in Table 1 indeed reflect a panel effect, since it is well-known that life satisfaction scores are correlated positively with income. The observation that individuals in the refreshment sample have, on average, a lower income than individuals in the more experienced sample is in line with the observation in the literature that people with lower income are likely to exit the panel (Kroh, 2011).

Moreover, as pointed out in Section 2.3, the refreshment samples are started well before the last round of data. To make experienced respondents more comparable with fresh respondents, Table 6 shows results from regression specifications described in Equation (1), restricting the refreshment samples to those individuals who will be in the panel for at least three more years. The results of this exercise seem to suggest that the discrepancy is even slightly higher. It thus seems that a selection effect would bias downward a panel effect rather than the reverse. Indeed, low life satisfaction scores are good predictors for future attrition (Kroh, 2011; Lipps, 2007).

Of course, in order for the latter robustness check to be convincing, the course of attrition over a sample's age should be similar for both the experienced and the refreshment samples. Attrition rates could be influenced by time-varying socioeconomic factors, or by factors related to the data collection (e.g. different interviewers, different

management, etc.). Large differences in attrition rates might then be an indication that the characteristics of the attritors across the experienced and the refreshment samples are not comparable. Figures 1, 2, and 3 show the attrition rate over the life cycle for the different samples for West Germany, East Germany, and Switzerland, respectively. An individual is regarded as an attritor at age n of a sample if the individual was a respondent at age 1, and if the individual answered the subjective well-being question strictly less than n times. The evolution of the attrition rates over a sample's life cycle within a panel are rather similar, concavely increasing. Attrition in the SHP is much higher than in the German SOEP, but this goes for both the 1999 sample as well as for the 2004 refreshment sample.

Finally, in order to investigate the size and dynamics of panel conditioning, regressions as of the form displayed in Equation (2) are run with N equal to 5. Hence, a subsample of individuals is selected who are re-interviewed in four or more consecutive years after the initial interview in which the question measuring subjective well-being was asked.⁵ Life satisfaction is regressed on five dummies taking the value of one if an individual is being interviewed for the first, second . . . , or fifth time respectively. Regression output is displayed in Table 7 for West Germany, East Germany, and Switzerland, respectively. The regression results show us that the panel effect is not entirely established between being interviewed for the first and second time but that it accumulates over the different survey rounds. In West Germany, there is a substantial panel effect from the second to the fifth interview, which is also statistically significant at any conventional significance level.⁶ There is an estimated panel effect of -0.13 and -0.11 during the second and third interview, respectively. There seems to be a panel effect of -0.09 during the fourth interview, and of -0.04 during the fifth interview. For East Germany, we see a similar pattern of negative panel effects recurring from interview to interview, although the null hypothesis that coefficients on subsequent interview indicators are equal can only be rejected with slightly higher P-values than for West Germany, ranging from 0.00 to 0.08. This is not surprising, since the sample sizes are much smaller.

⁵In the SHP, the first interview for an individual does not necessarily equal the interview in which he was first asked about his life satisfaction, as a question on subjective well-being was only introduced in the second round of the panel.

⁶To be clear, a panel effect for interview n is calculated by subtracting the coefficient on the dummies for interview $n - 1$ and n .

The pattern for Switzerland is in line with that of Germany, but no panel effect is measured any longer after the third interview, and a panel effect is only statistically significant (be it at any conventional level) for the second interview.

The coefficient on the dummy for being interviewed for the fifth time can be interpreted as the negative of a residual panel effect, that is, which will be established over the interviews after the fifth has taken place. The coefficient on this dummy for Switzerland is nearly equal to 0, while for West and East Germany, this coefficient still has a substantial magnitude of 0.20 and 0.27, respectively. The reason why the cumulation path of panel effects for Switzerland is much shorter than for Germany is rather speculative, but one should note that in the SHP, 90% of people answering the well-being question for the first time were asked in either 2000 or 2004. Moreover, one should keep in mind that, contrary to the path of panel effects from the first to the fifth interviews, the estimate for a residual panel effect might be slightly clouded by panel attrition. Indeed, some respondents will no longer be interviewed after the fifth interview, while others will remain in the panel for many years.

3.2 A Comparison between Time Trends in Panel Data and Repeated Cross-Sections

Comparing time trends between panel data and repeated cross-sectional data within the same country is another, though less ideal strategy to identify the existence of a panel effect. First, the design of the questionnaires used to construct the panel datasets and the repeated cross-sectional datasets are different, as well as the question wording and the answering scales for the well-being indicators. Moreover, comparisons are made difficult because repeated cross-sectional data are not always available with yearly intervals, and subtle changes in survey design in both panel data and repeated cross-sections might lead to artificial results (Conti and Pudney, 2011), and it seems difficult to disentangle a panel effect from an attrition effect. Nevertheless, it may be useful to check whether the results from this exercise are in line with the results described in the previous subsection.

Figure 4 shows us trends of life satisfaction from the German SOEP for the West German sample, for men and women separately. During the first five years of the panel, we see an overall decline in well-being from 7.4 to 7.0. In the next three years, following reunification, we see an upsurge in well-being, after which there is an almost steady

decline until 1997, when well-being reaches a dip of 7. During the next four years (a period of economic recovery and when two refreshment samples are being introduced), well-being increased again up to a score of 7.3 in 2001. From 2002 to 2010, the pattern is a bit more irregular but overall downward sloping (with a dip in 2004). The German SOEP data seem to suggest, for West Germany, that well-being over time is overall very slightly downward sloping. It seems, over these 27 years, that the well-being for both men and women has undergone a very similar pattern. Moreover, the two groups report, on average, the same level of life satisfaction at both the beginning and the end of the period analysed.

Figure 5 shows us trends of life satisfaction for East Germany from 1990 onwards. Life satisfaction in East Germany is considerably lower than in West Germany. This gap narrows near the end of the sample but is still very pronounced, even 20 years after reunification. The huge drop in life satisfaction between 1990 and 1991 followed by a recovery is remarkable and is consistent with results found for East Germany and other transition countries by Sanfey and Teksoz (2007), Easterlin (2009), and Easterlin and Plagnol (2009). From 1991 onwards, the overall pattern for East Germany is increasing. This discrepancy with West German data might be interpreted as evidence of a reduction in relative deprivation (in many ways) compared with neighbouring countries.

If we now turn to German data from the WVS, comparison of means between outcomes for 1997 and 2006 reveals no downward trend. If anything, average well-being is even slightly higher in the latter year than in the former (2.973 versus 2.966). Figure 6 shows us trends in standardized life satisfaction scores from SOEP data and Eurobarometer data for West Germany over the period 1984 to 1998. As subjective well-being data is missing in the Eurobarometer Survey for 1996, a score for this year is imputed by averaging scores of 1995 and 1997. Generally, well-being in the Eurobarometer Survey is upward sloping and higher in 1995 than in 1984. The sudden dip at the end of the period is remarkable. Stevenson and Wolfers (2009) have been puzzling over these West German results as well. When looking at each country separately in the Eurobarometer Survey, they find West Germany the exception wherein life satisfaction was lower at the end of the period than at the beginning. One should note, however, that the context in which the trend question is being asked can vary across different survey rounds. For example, the 1995 data are drawn from a questionnaire conducted during the period April-May, which placed the life satisfaction second in sequence (after asking about one's nationality). In 1997, the life satisfaction question was the thirty-sixth question in

a questionnaire conducted during the period March-April. As earlier research suggests, placement of a question in a questionnaire might influence reported well-being scores (Conti and Pudney, 2011). Moreover, conducting a survey earlier in the year might influence results for reported well-being, for example, due to a lower amount of sunshine (Güven, 2012). Figure 7 shows us the results from the same data sources for East Germany for the period 1990 to 1998. The comparison between both datasets suggests that the very significant drop in life satisfaction between 1990 and 1991 shown by German SOEP data is most likely not entirely due to a panel effect.

Figure 8 shows us trends in GHQ scores for the whole sample and men and women separately from 1991 to 2008 onwards. A GHQ score of 0 means ‘seriously depressed’ and 12 means ‘no depressive symptoms’. It is remarkable that GHQ scores for women are considerably lower than those for men.⁷ This reminds us that a depression index is conceptually different from a life satisfaction score. The GHQ scores show an overall slightly downward sloping pattern, with peaks in 1999 and a dip in the first years of the millennium. In 2008, there is a huge jump in GHQ scores, for both men and women, of more than one point. This result for 2008 is quite remarkable and puzzling. The positioning of the GHQ questions does not seem to have changed in the questionnaire, and the increase is not caused by new individuals entering the panel. Figure 9 shows us trends in life satisfaction from 1996 onwards. As no data are available for 2001, the value for 2001 has been imputed by averaging the mean of scores in 2000 and 2002. Women and men report more or less the same life satisfaction scores. Again, there is a peak in life satisfaction in 2008. According to the British WVS data, on average, a well-being score of 3.21 is reported in 1998, increasing to 3.42 in 2006. This increase is statistically significant at any significance level. Figure 10 shows us standardized GHQ scores from the BHPS and standardized life satisfaction scores from the Eurobarometer Survey for the period 1991 to 1998. The standardized GHQ scores from the BHPS (aggregated for men and women) show an overall downward sloping trend over the examined period, while there is no such observation for the life satisfaction scores in the Eurobarometer Survey, which show no clear trend.⁸

⁷This finding is not just limited to Britain, as unreported results from the Panel Study of Belgian Households suggest as well.

⁸Unfortunately, single-item life satisfaction scores from the BHPS cannot be used in this exercise, since general life satisfaction in the BHPS is only measured from 1996 onwards and since life satisfaction data from the Eurobarometer survey is missing for 1996.

Life satisfaction time trends derived from the SHP are shown in Figure 11 for the whole population, and for men and women separately, for the years 2000 to 2010. The life satisfaction measure in the SHP has a 0-10 range, as does the life satisfaction measure in the German SOEP. This means that we can quite easily compare the averages of the scores. Swiss respondents rate their satisfaction with life, on average, much higher than Germans do. In West Germany, average life satisfaction over time ranges from 6.9 to 7.4, while in Switzerland, the average scores range from 7.9 to 8.2. Of course, even though the questions are identical, the surveys' designs are not the same. For example, while the German SOEP gathers data via face-to-face interviews, the SHP interviews are conducted by telephone. Anchoring effects might play a role as well. On average, there seems to be a slight decrease over time. The time span is rather short, so observing convergence or divergence in male and female happiness levels remains difficult. Female happiness is slightly higher than male happiness in the beginning of the period but is equal to male happiness by the end.

As for Swiss WVS data, average well-being scores equal 3.29 in 1989, 3.34 in 1996, and 3.35 in 2007. The increase between 1989 and 1996 is statistically significant at the 7% significance level, the increase over the period 1996 to 2007 is not economically significant, nor statistically significantly different from 0. The increase between 1989 and 2007, however, is statistically significant at the 4% significance level and seems to be of a substantial magnitude as well (0.05), given the limited scale of the variable and the relatively short time span. Unfortunately, Eurobarometer data for Switzerland are not available.

It thus seems that cross-sectional data show us a flat or slightly increasing trend in life satisfaction over time, while panel data show us an overall downward trend (except for East Germany post-reunification). To check whether this downward trend is genuine and not some random noise, one can fit the well-being data with just a linear time trend instead of a full set of time dummies, and check whether the overall time trend is negative and significantly different from 0.

Fitting the life satisfaction data for West German SOEP (using weights to correct for panel ageing) gives us a negative time trend of 0.019 points per year between 1984 and 1998 (the period for which comparison with Eurobarometer data is possible), and of 0.018 points from 1999 to 2010. For Britain, GHQ scores go down by, on average, 0.06 point per year between 1991 and 1998. Again, this trend is significantly different from 0 at all significance levels. The trend after 1998 is slightly negative, although not

significant, and only after ignoring the exceptional outlier of 2008. Also for Switzerland, a significantly negative trend in life satisfaction data of 0.016 points, can be identified in the SHP for the period 2000-2010.

These results seem quite encouraging, since the different factors which might cloud the results are likely to do this in the direction opposite to the identified trends. First, as shown above, panel attrition is likely to lead to an increase in average life satisfaction over time, since less happy people are more likely to exit the panel. Second, while on average, the panel grows more experienced over time, new refreshers (either natural refreshers or induced refreshers) are entering the panel, which, by construction, will partly mitigate the panel effect.⁹

4 Conclusion

This paper has investigated whether subjective well-being data are subject to panel conditioning, that is, whether answers on questions measuring subjective well-being depend on having participated in previous rounds of the panel. The analysis has used panel data as well as repeated cross-sectional data for three European countries.

The patterns of the life satisfaction measures in the different panel datasets show that in most cases, scores are slightly lower in the last rounds of data than they are in the first rounds, even when weights are used to correct for panel ageing. Patterns in life satisfaction differ little between men and women (likely because of the relatively short time span of the data), but the data from the BHPS show that women are more likely to report depressive symptoms than men are. When comparing panel data results with repeated cross-sectional results from the WVS and the Eurobarometer Survey, one observes some discrepancies. While panel data show for different countries (apart from East Germany) an overall downward trend, repeated cross-sections for these countries indicate an increase in well-being over time, and this comparison offers some first (mild) evidence for a panel effect. More convincing evidence seems to emerge when comparing refreshment samples during their first year with the data in the more experienced sample: subjective well-being scores are substantially higher in the refreshment sample than in the more experienced sample in the corresponding calendar year. Most important, these

⁹Looking at a balanced panel is not very useful since this would imply that young ages will not be present at the end of the time period, and old ages will be absent in the first periods.

results do not seem to be driven by panel attrition, and it seems that panel effects accumulate across subsequent interviews.

Strong evidence concerning the existence of a panel effect has important implications for economists and policy makers using subjective well-being data. First, the results can guide the increasing literature on trends in well-being, as they suggest that repeated cross-sections and pseudo panel data might be favoured over genuine panel data in studying the latter phenomenon. Second, they might help fine tune policy evaluation design. For example, oversampling a minority in a certain year of a household panel to compare its well-being scores with those of the experienced subsample might lead us to wrong conclusions, even if attrition in the older sample is negligible. Finally, the identification of a substantial panel effect in subjective well-being data might help draw attention to this phenomenon in other areas of economics.

References

Becchetti, L., A. Pelloni and F. Rossetti (2008) Relational Goods, Sociability, and Happiness. *Kyklos* 61, 343-363.

Beegle, K., H. Himelein and M. Ravallion (2012) Frame-of-reference Bias in Subjective Welfare. *Journal of Economic Behavior and Organization* 81, 556-570.

Bertrand, M. and S. Mullainathan (2001) Do People Mean What They Say? Implications for Subjective Survey Data. *American Economic Review* 91, 67-73.

Blanchflower, D. and A. Oswald (2004) Well-being over Time in Britain and the USA. *Journal of Public Economics* 88, 1359-1386.

Blanchflower, D. and A. Oswald (2008) Is Well-being U-shaped over the Life Cycle? *Social Science and Medicine* 66, 1733-1749.

Blanchflower, D., A. Oswald and B. Van Landeghem (2009) Imitative Obesity and Relative Utility. *Journal of the European Economic Association* 7, 528-538.

Bonsang, E. and A. van Soest (2012) Satisfaction with Job and Income among Older Individuals across European Countries. *Social Indicators Research* 105, 227-254.

Clark, A., E. Diener, Y. Georgellis and R. Lucas (2007) Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis. *Economic Journal* 118, 222-243.

Clark, A. and A. Oswald (2006) The Curved Relationship between Subjective Well-being and Age. *PSE Working Papers* 2006-29.

Clark, A. and C. Senik (2010) Will GDP Growth Increase Happiness in Developing Countries? *PSE Working Papers* 2010-43.

Conti, G. and S. Pudney (2011) Survey Design and the Analysis of Satisfaction. *Review of Economics and Statistics* 93, 1087-1093.

Das, J., A. van Soest and V. Toepoel (2011) Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys. *Sociological Methods and Research* 40, 32-56.

Deaton, A. (2008) Income, Health, and Well-being around the World: Evidence from the Gallup World Poll. *Journal of Economic Perspectives* 22, 53-72.

de Grip, A., M. Lindeboom and R. Montizaan (2012) Shattered Dreams: The Effects of Changing the Pension System Late in the Game. *The Economic Journal* 122, 1-25.

De Neve, J.-M., N. Christakis, J. Fowler and B. Frey (2010) Genes, Economics, and Happiness. *CESIFO Working Paper Series* No. 2946.

De Neve, J.-M. and A. Oswald (2012) Happiness Pays: Using Sibling Fixed Effects to Estimate the Influence of Psychological Well-Being on Later Income. *Royal Economic Society 2012 Annual Conference paper*, University of Cambridge.

Di Tella, R. and R. MacCulloch (2008) Gross National Happiness as an Answer to the Easterlin Paradox? *Journal of Development Economics* 86, 22-42.

Di Tella, R., R. MacCulloch and A. Oswald (2001) Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review* 91, 335-341.

Di Tella, R., R. MacCulloch and A. Oswald (2003) The Macroeconomics of Happiness. *Review of Economics and Statistics* 85, 809-827.

Easterlin, R. (1974) Does Economic Growth Improve the Human Lot? In Paul A. David and Melvin W. Reder, eds., *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, New York: Academic Press, Inc.

Easterlin, R. (1995) Will Raising the Incomes of All Increase the Happiness of All? *Journal of Economic Behavior and Organization* 27, 35-48.

Easterlin, R. (2001) Income and Happiness: Towards a Unified Theory. *Economic Journal* 111, 465-484.

Easterlin, R. (2006) Life Cycle Happiness and Its Sources: Intersections of Psychology, Economics, and Demography. *Journal of Economic Psychology* 27, 463-482.

Easterlin, R. (2009) Lost in Transition: Life Satisfaction on the Road to Capitalism. *Journal of Economic Behavior and Organization* 71, 130-145.

Easterlin, R., L. Angelescu-McVey, M. Switek, O. Sawangfa and J. Smith-Zweig (2010) The Happiness-Income Paradox Revisited. *Proceedings of the National Academy of Sciences* 107, 22463-22468.

Easterlin, R. and A. Plagnol (2009) Life Satisfaction and Economic Conditions in East and West Germany Pre- and Post-Unification. *Journal of Economic Behavior and Organization* 68, 433-444.

Ferrer-i-Carbonell, A. and P. Frijters (2004) How Important Is Methodology for the Estimates of the Determinants of Happiness? *Economic Journal* 114, 641-659.

Fischer, J. (2009) Happiness and Age Cycles - Return to Start. . . . *MPRA Paper* No. 15249, University of Munich.

Frijters, P. and T. Beaton (2012) The Mystery of the U-shaped Relationship between Happiness and Age. *Journal of Economic Behavior and Organization* 82, 525-542.

Gerstorff, D., N. Ram, G. Mayraz, M. Hidajat, U. Lindenberger, G. Wagner and J. Schupp (2010) Late-life Decline in Well-being across Adulthood in Germany, the United Kingdom, and the United States: Something Is Seriously Wrong at the End of Life. *Psychology and Aging* 25, 477-485.

Güven, C. (2012) Reversing the Question: Does Happiness Affect Consumption and Savings Behavior? *Journal of Economic Psychology* 33, 701-717.

Gwozdz, W. and a. Sousa-Poza (2010) Ageing, Health and Life Satisfaction of the Oldest Old: An Analysis for Germany. *Social Indicators Research* 97, 397-417.

Ifcher, J. and O. Zarghamee (2011) Happiness and Time Preference: The Effect of Positive Affect in a Random-Assignment Experiment. *American Economic Review* 10, 3109-3129.

Kapteyn, A., J. Smith and A. van Soest (2011) Are Americans Really Less Happy with Their Incomes? *Discussion Paper* 2011-059, Tilburg University, Center for Economic Research.

Kassenboehmer, S. and J. Haisken-DeNew (2012) Heresy or Enlightenment? The Well-being Age U-shape Effect Is Flat. *Economics Letters* 117, 235-238.

Keisuke, H., G. Imbens, G. Ridder and D. Rubin (2001) Combining Panel Data Sets with Attrition and Refreshment Samples. *Econometrica* 69, 1645-1659.

Kroh, M. (2011) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2010). *Data Documentation* 59, Deutsches Institut für Wirtschaftsforschung.

Krueger, A. and D. Schkade (2008) The Reliability of Subjective Well-being Measures. *Journal of Public Economics* 92, 1833-1845.

Lipps, O. (2007) Attrition in the Swiss Household Panel. *Methoden - Daten - Analysen* 1, 45-68.

Luttmer, E. (2005) Neighbours as Negatives: Relative Earnings and Well-being. *Quarterly Journal of Economics* 120, 963-1002.

McBride, M. (2010) Money, Happiness, and Aspirations: An Experimental Study. *Journal of Economic Behavior and Organization* 74, 262-276.

Metcalf, R., N. Powdthavee and P. Dolan (2011) Destruction and Distress: Using a Quasi-experiment to Show the Effect of the September 11 Attacks on Mental Well-being in the United Kingdom. *Economic Journal* 121, F81-F103.

Oswald, A. (2008) On the Curvature of the Reporting Function from Objective Reality to Subjective Feelings. *Economics Letters* 100, 369-72.

Oswald, A. and N. Powdthavee (2008a) Death, Happiness, and the Calculation of Compensatory Damages. *Journal of Legal Studies* 37, S217-S252.

Oswald, A. and N. Powdthavee (2008b) Does Happiness Adapt? A Longitudinal Study of Disability with Implications for Economists and Judges. *Journal of Public Economics* 92, 1061-1077.

Oswald, a., E. Proto and D. Sgroi (2012) Happiness and Productivity. *Working Paper*, University of Warwick.

Oswald, A. and S. Wu (2010) Objective Confirmation of Subjective Measures of Human Well-being: Evidence from the U.S.A. *Science* 327, 576-580.

Oswald, A. and S. Wu (2011) Well-being across America. *review of Economics and Statistics* 93, 1118-1134.

Powdthavee, N. (2008) Putting a Price Tag on Friends, Relatives, and Neighbours: Using Surveys of Life Satisfaction to Value Social Relationships. *Journal of Socio-Economics* 37, 1459-1480.

Proto, E., D. Sgroi and A. Oswald (2012) Are Happiness and Productivity Lower among University Students with Newly-Divorced Parents? An Experimental Approach. *Experimental Economics* 15, 1-23.

Ravallion, M. and M. Lokshin (2010) Who Cares About Relative Deprivation? *Journal of Economic Behavior and Organization* 73, 171-185.

Sacks, D., B. Stevenson and J. Wolfers (2010) Subjective Well-being, Income, Economic Development, and Growth. *NBER Working Paper* No. 16441, National Bureau of Economic Research.

Sanfey, P. and U. Teksoz (2007) Does Transition Make You Happy? *Economics of Transition* 15, 707-731.

Scholz, E. and H. Schmitt (2001) The Mannheim Eurobarometer Trend File 1970 - 1999 (Codebook and Unweighted Frequency Distributions). *Report*, Mannheim.

Senik, C. (2004) When Information Dominates Comparison: Learning From Russian Subjective Panel Data. *Journal of Public Economics* 88, 2099-2123.

Senik, C. (2008a) Ambition and Jealousy: Income Interaction in the “Old” (Europe Versus the “New” Europe and the United States. *Economica* 75, 495-513.

Senik, C. (2008b) Is Man Doomed to Progress? *Journal of Economic Behavior and Organization* 68, 140-152.

Senik, C. (2009) Direct Evidence on Income Comparisons and Their Welfare Effects. *Journal of Economic Behavior and Organization* 72, 408-424.

Sgroi, d., E. Proto, A. Oswald and A. Dobson (2010) Priming and the Reliability of Subjective Well-being Measures. *Warwick Economic Research Papers* No. 935.

Sharpe, J. and D. Gilbert (1998) Effects of Repeated Administration of the Beck Depression Inventory and Other Measures of Negative Mood States. *Personal Individual Differences* 24, 457-463.

Stevenson, B. and J. Wolfers (2008) Economic Growth and Subjective Well-being: Reassessing the Easterlin Paradox. *NBER Working Papers* No. 14282, National Bureau of Economic Research.

Stevenson, B. and J. Wolfers (2009) The Paradox of Declining Female Happiness. *American Economic Journal: Economic Policy* 1, 190-225.

Taylor, M., J. Brice, N. Buck and E. Prentice-Lane (2010) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.

Toepoel, V., J. Das and A. van Soest (2009) Relating Question Type to Panel Conditioning: A Comparison between Trained and Fresh Respondents. *Survey Research Methods* 3, 73-80.

Van Landeghem, B. (2012) A Test for the Convexity of Human Well-being over the Life Cycle: Evidence from a 20-Year Panel. *Journal of Economic Behavior and Organization* 81, 571-582.

Voorpostel, M., R. Tillmann, F., Lebert, B. Weaver, U. Kuhn, O., Lipps, V.-A. Ryser, F. Schmid, M. Rothenbuehler and B. Wernli (2011) *Swiss Household Panel Userguide (1999-2010), Wave 12, October 2011*. Lausanne: FORS.

Wagner, G., J. Frick and J. Schupp (2007) The German Socio-economic Panel Study (SOEP) - Scope, Evolution, and Enhancements. *Schmollers Jahrbuch* 127, 139-169.

Wolfers, J. (2003) Is Business Cycle Volatility Costly? Evidence from Surveys of Subjective Well-being. *International Finance* 6, 1-26.

World Values Survey Association (www.worldvaluessurvey.org). *WORLD VALUES SURVEY 1981-2008 OFFICIAL AGGREGATE v.20090901, 2009*. Aggregate File Producer: ASEP/JDS, Madrid.

Table 1: Differences in Well-being between Refreshment Samples and Experienced Samples

	West Germany	East Germany	Switzerland
$E[W_{1998}^{re}] - E[W_{1998}^{exp}]$	0.55	0.48	
$E[W_{2000}^{re}] - E[W_{2000}^{exp}]$	0.37	0.27	
$E[W_{2004}^{re}] - E[W_{2004}^{exp}]$			0.16
$E[W_{2006}^{re}] - E[W_{2006}^{exp}]$	0.43	0.64	

Source: German Socioeconomic Panel and Swiss Household Panel.

W = well-being, subscripts denote the survey year, superscripts RE and EXP denote ‘refreshment sample’ and ‘experienced sample’, respectively. All differences turn out to be significantly different from 0 at any significance level.

Table 2: Differences in Well-being between Refreshment Samples and Experienced Samples: Male Subsample

	West Germany	East Germany	Switzerland
$E[W_{1998}^{re}] - E[W_{1998}^{exp}]$	0.58	0.47	
$E[W_{2000}^{re}] - E[W_{2000}^{exp}]$	0.34	0.30	
$E[W_{2004}^{re}] - E[W_{2004}^{exp}]$			0.11
$E[W_{2006}^{re}] - E[W_{2006}^{exp}]$	0.39	0.56	

Source: German Socioeconomic Panel and Swiss Household Panel.

W = well-being, subscripts denote the survey year, superscripts RE and EXP denote ‘refreshment sample’ and ‘experienced sample’, respectively. All differences turn out to be significantly different from 0 at any significance level, except for the Swiss subsample, where the difference of 0.11 is only statistically different from 0 at the 3% significance level.

Table 3: Differences in Well-being between Refreshment Samples and Experienced Samples: Female Subsample

	West Germany	East Germany	Switzerland
$E[W_{1998}^{re}] - E[W_{1998}^{exp}]$	0.52	0.49	
$E[W_{2000}^{re}] - E[W_{2000}^{exp}]$	0.39	0.25	
$E[W_{2004}^{re}] - E[W_{2004}^{exp}]$			0.20
$E[W_{2006}^{re}] - E[W_{2006}^{exp}]$	0.46	0.72	

Source: German Socioeconomic Panel and Swiss Household Panel.

W = well-being, subscripts denote the survey year, superscripts RE and EXP denote ‘refreshment sample’ and ‘experienced sample’, respectively. All differences turn out to be significantly different from 0 at any significance level.

Table 4: Exploring the Determinants of the Propensity to Belong to a Refreshment Sample: Results for West Germany, East Germany, and Switzerland

	1998		2000		2006		2004
	West Germany	East Germany	West	East	West	East	Switzerland
log real hh. income	-0.043*** (0.012)	-0.027** (0.011)	-0.094*** (0.012)	-0.011 (0.019)	-0.002 (0.011)	-0.010 (0.011)	-0.051*** (0.012)
disabled	-0.056*** (0.015)	0.016 (0.018)	-0.070*** (0.017)	-0.005 (0.026)	-0.008 (0.012)	-0.014 (0.015)	0.020 (0.013)
household size	-0.000 (0.006)	-0.015** (0.006)	0.033*** (0.006)	0.006 (0.009)	0.009* (0.005)	-0.008 (0.006)	-0.030*** (0.006)
low education	-0.008 (0.017)	0.167*** (0.040)	-0.074*** (0.018)	0.094*** (0.029)	-0.033* (0.019)	0.007 (0.027)	0.053*** (0.020)
medium education	-0.046*** (0.015)	0.055*** (0.011)	0.017 (0.014)	0.234*** (0.017)	-0.056*** (0.014)	-0.050*** (0.017)	0.006 (0.021)
unemployed	-0.066*** (0.024)	0.002 (0.013)	-0.018 (0.032)	-0.016 (0.022)	0.006 (0.020)	0.006 (0.015)	0.043 (0.043)
single	0.013 (0.017)	0.040*** (0.013)	0.021 (0.017)	-0.062*** (0.024)	-0.017 (0.015)	0.027* (0.015)	-0.047** (0.019)
widowed	-0.045** (0.019)	-0.037*** (0.013)	-0.046** (0.022)	-0.020 (0.031)	0.033 (0.021)	-0.004 (0.021)	-0.012 (0.032)
divorced	-0.043** (0.020)	-0.008 (0.015)	-0.047** (0.022)	0.048* (0.027)	-0.002 (0.016)	-0.012 (0.015)	-0.037* (0.022)
female	-0.019* (0.012)	-0.003 (0.009)	-0.004 (0.011)	0.024* (0.014)	-0.023** (0.011)	0.001 (0.011)	-0.040*** (0.012)
Observations	5,049	3,569	9,520	5,169	8,439	4,581	6,707

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The table shows marginal probabilities, computed at the average value of independent variables, derived from probit equations. Regressions are run for each survey year in which a refreshment sample was started. The dependent variable is dichotomous and takes the value 1 whenever an individual belongs to the refreshment sample started in that particular year.

Table 5: Differences in Real Household Income between Refreshment Samples and Experienced Samples (Expressed in %)

	West Germany	East Germany	Switzerland
$E[Y_{1998}^{re}] - E[Y_{1998}^{exp}]$	-9.6***	-16.7***	
$E[Y_{2000}^{re}] - E[Y_{2000}^{exp}]$	-6.8***	-6.5***	
$E[Y_{2004}^{re}] - E[Y_{2004}^{exp}]$			-6.2***
$E[Y_{2006}^{re}] - E[Y_{2006}^{exp}]$	-0.2	-0.4	

Source: German Socioeconomic Panel and Swiss Household Panel.

Y = real household income, subscripts denote the survey year, superscripts RE and EXP denote ‘refreshment sample’ and ‘experienced sample’, respectively. One to three asterisks denote significance at the 10%, 5%, and 1% significance levels, respectively.

Table 6: Differences in Well-being between Refreshment Samples and Experienced Samples: Addressing Attrition Bias

	West Germany	East Germany	Switzerland
$E[W_{1998}^{re}] - E[W_{1998}^{exp}]$	0.65	0.47	
$E[W_{2000}^{re}] - E[W_{2000}^{exp}]$	0.48	0.32	
$E[W_{2004}^{re}] - E[W_{2004}^{exp}]$			0.21
$E[W_{2006}^{re}] - E[W_{2006}^{exp}]$	0.49	0.63	

Source: German Socioeconomic Panel and Swiss Household Panel.

WB = well-being, subscripts denote the survey year, superscripts RE and EXP denote ‘refreshment sample’ and ‘experienced sample’, respectively. All differences turn out to be significantly different from 0 at any significance level. As for the refreshment sample, only individuals who will remain in the panel for at least three more years are considered.

Table 7: The Path of Panel Effects over Interviews: Results for West Germany, East Germany, and Switzerland

VARIABLES	West Germany sat	East Germany sat	Switzerland sat
1st Interview	0.590*** (0.019)	0.691*** (0.038)	0.234*** (0.055)
2nd Interview	0.456*** (0.018)	0.501*** (0.036)	0.103** (0.043)
3rd Interview	0.345*** (0.017)	0.433*** (0.034)	0.056 (0.039)
4th Interview	0.246*** (0.017)	0.324*** (0.032)	0.032 (0.037)
5th Interview	0.202*** (0.016)	0.273*** (0.032)	0.042 (0.032)
Time Dummies	YES	YES	YES
Constant	6.864*** (0.033)	5.506*** (0.055)	8.012*** (0.065)
Observations	221,167	79,091	50,217
R-squared	0.014	0.017	0.005

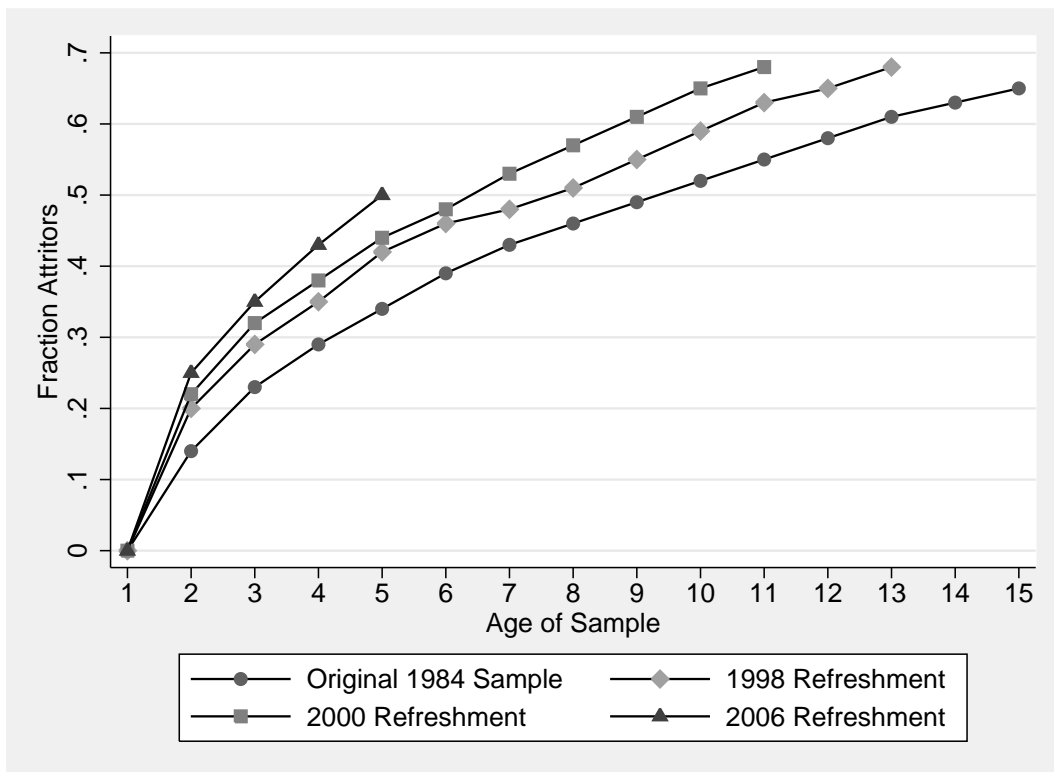
Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: German Socioeconomic Panel and Swiss Household Panel.

Regressions are run on individuals who have answered at least the subjective questions in the four survey rounds following the interview in which they were first asked the subjective well-being question.

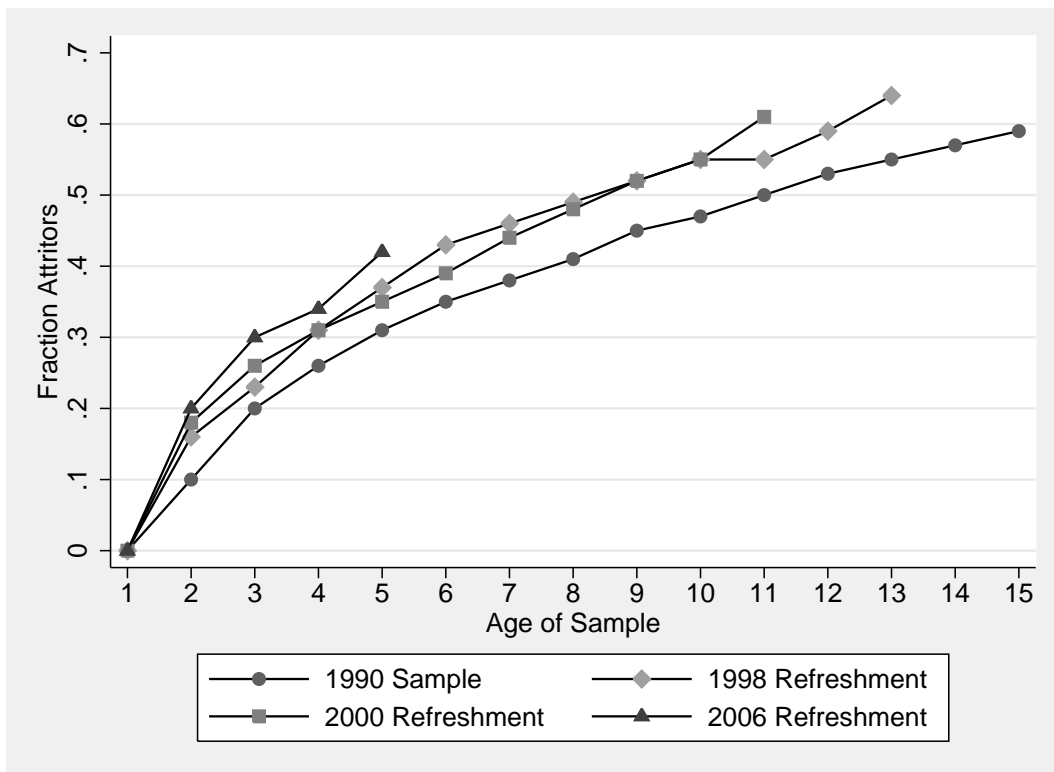
Figure 1: The Attrition Rate over a Sample's Life Cycle: West Germany



Source: German Socioeconomic Panel.

An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n .

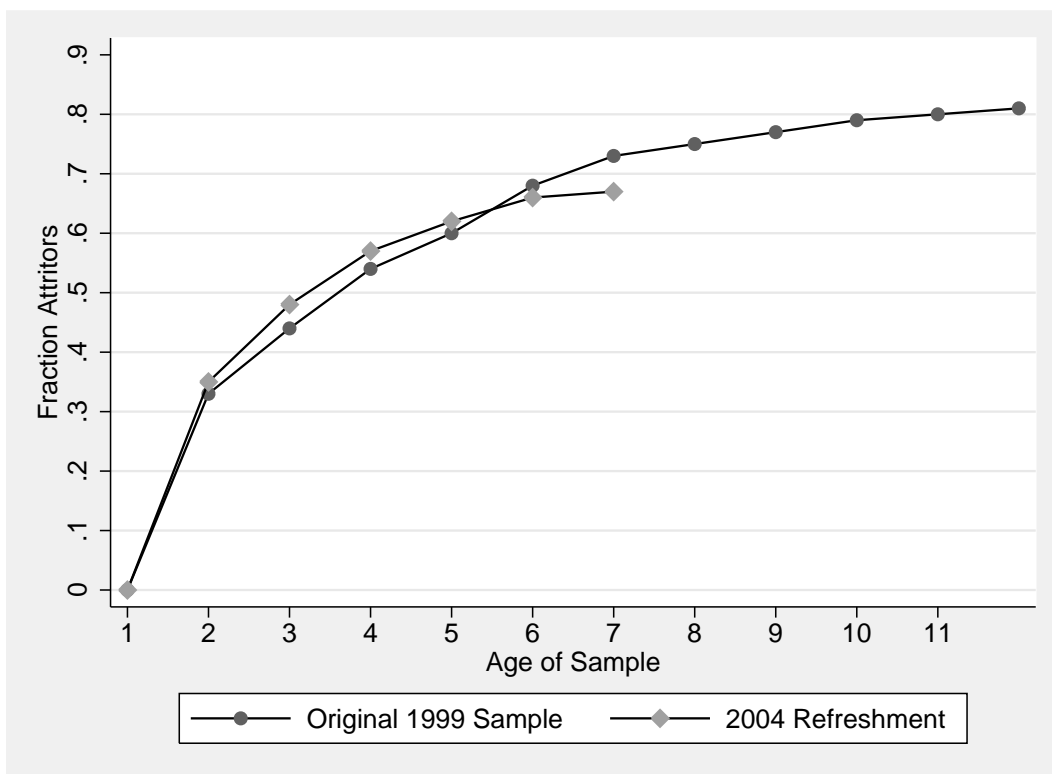
Figure 2: The Attrition Rate over a Sample's Life Cycle: East Germany



Source: German Socioeconomic Panel

An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n .

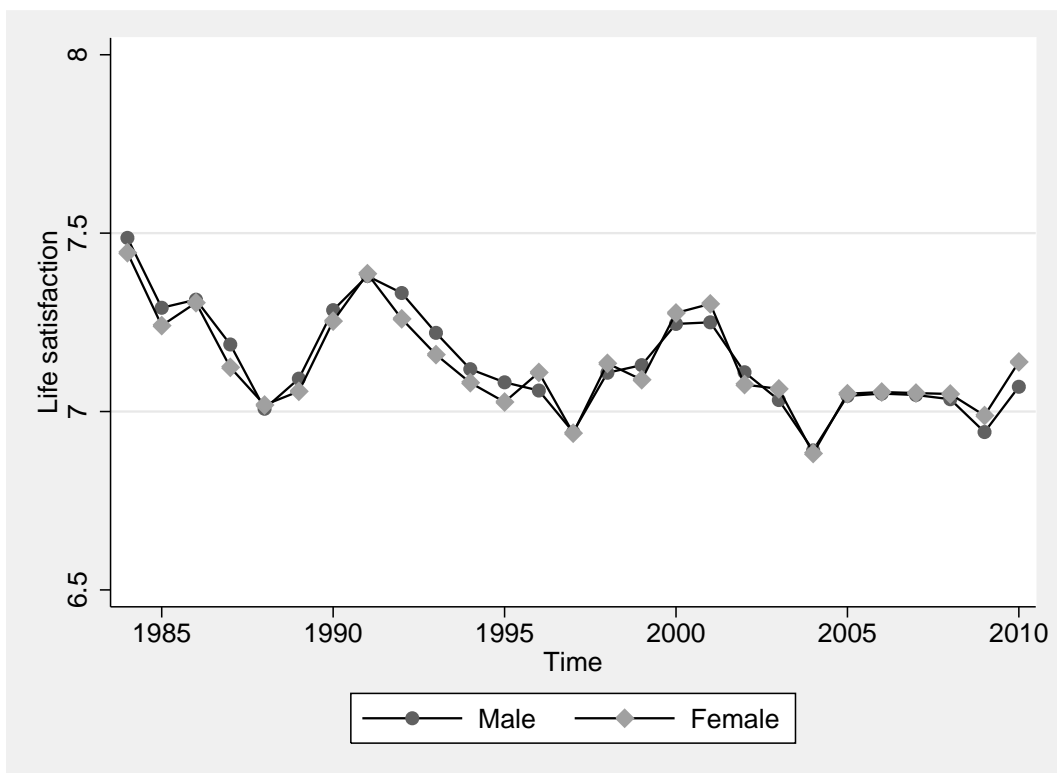
Figure 3: The Attrition Rate over a Sample's Life Cycle: Switzerland



Source: Swiss Household Panel

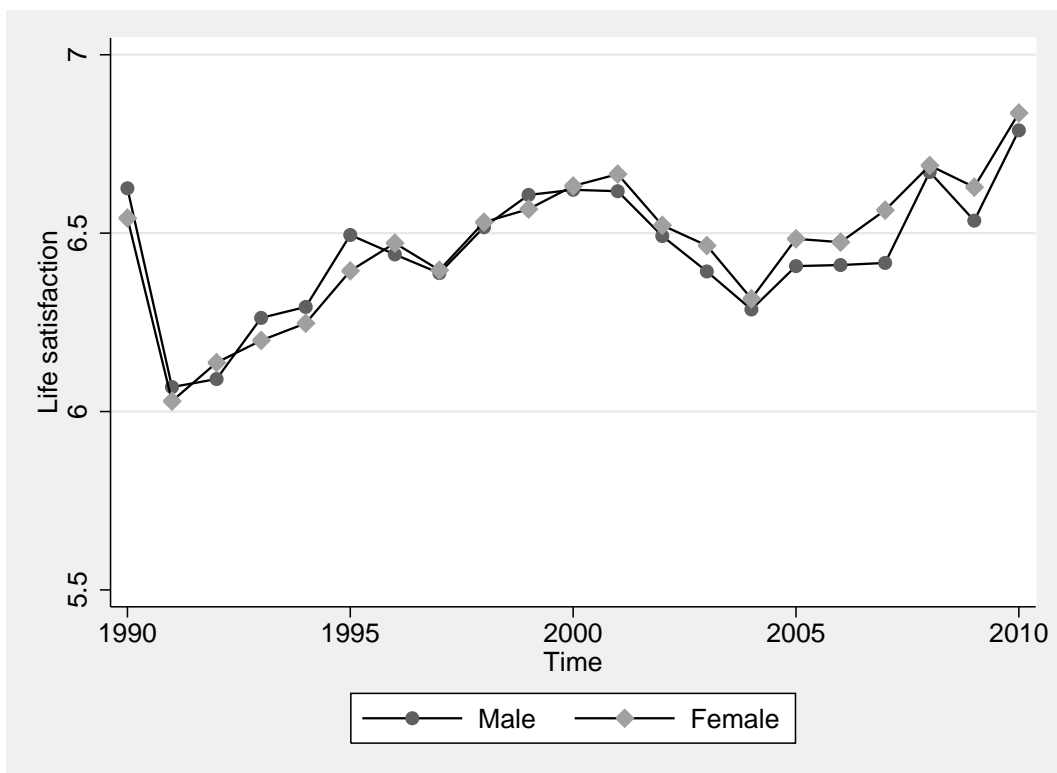
An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n .

Figure 4: Trends of Life Satisfaction over Time: West Germany



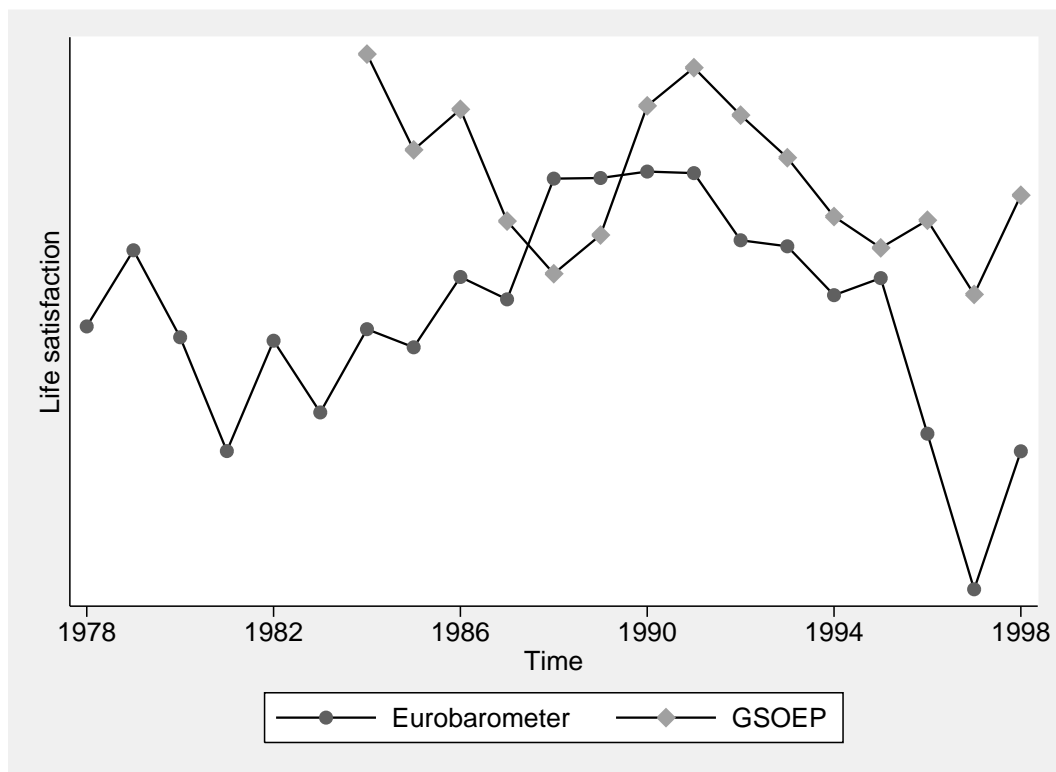
Source: German Socioeconomic Panel.

Figure 5: Trends of Life Satisfaction over Time: East Germany



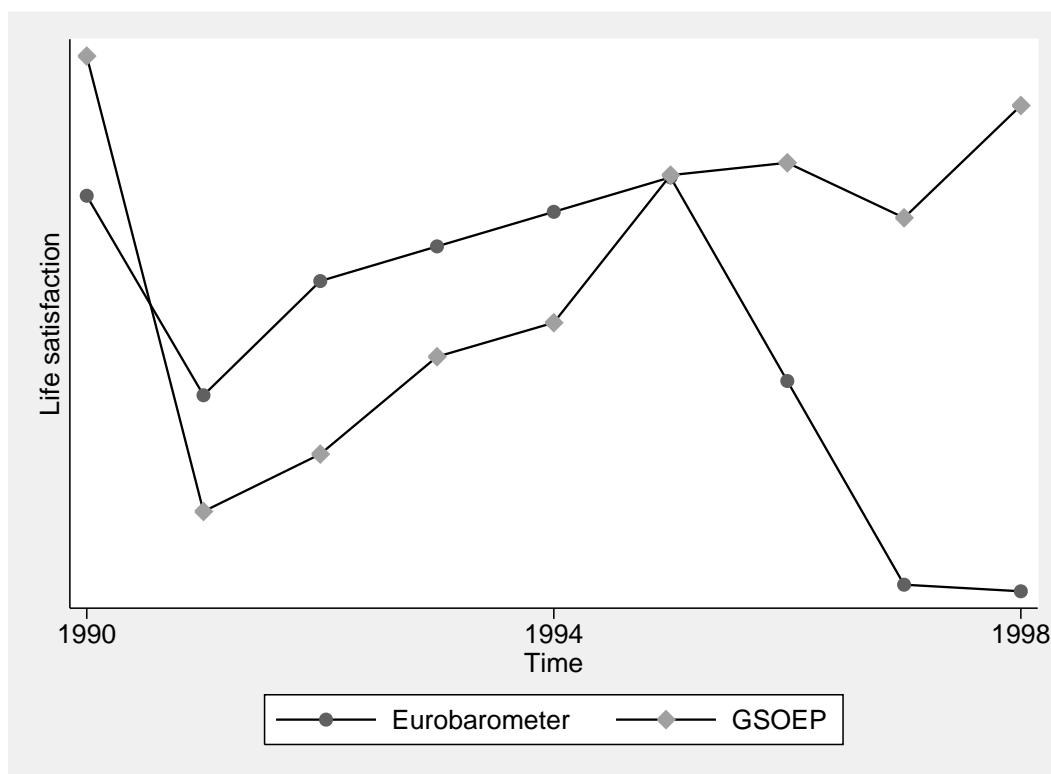
Source: German Socioeconomic Panel.

Figure 6: Comparison between Trends in Standardized Life Satisfaction in the German SOEP and the Eurobarometer Survey: West Germany



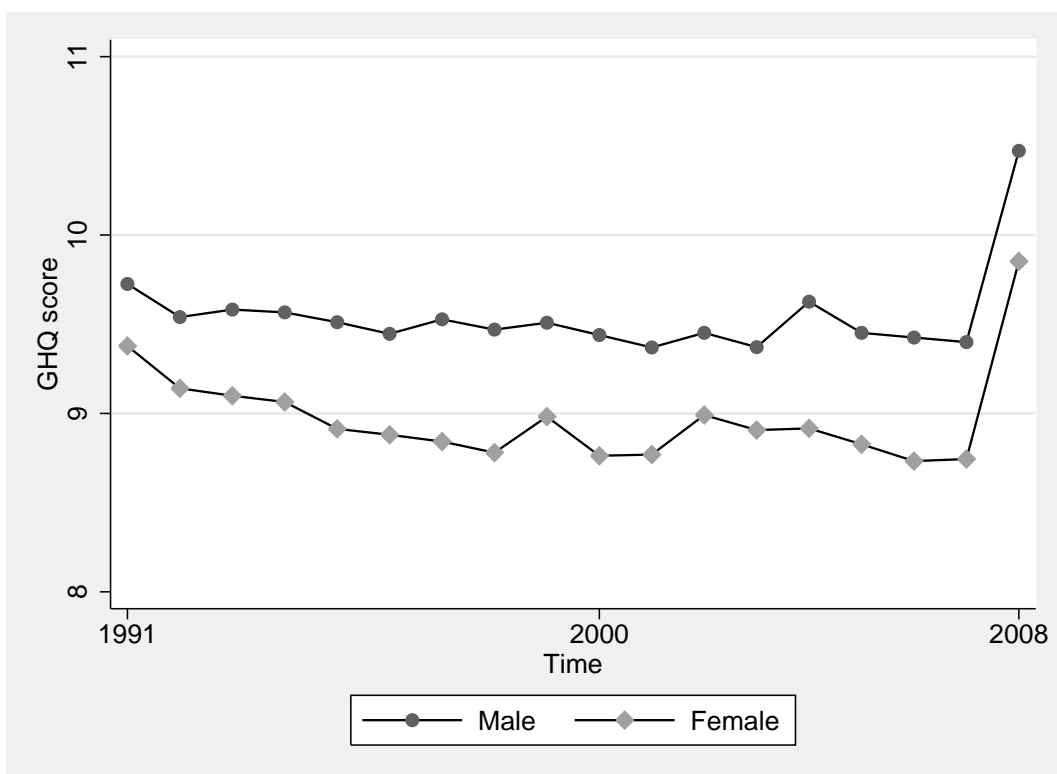
Source: German Socioeconomic Panel and Eurobarometer Survey. For both data sources, life satisfaction has been standardized with mean 0 and standard error 1. Hence, the graphs are only meant for illustrative purposes, and they should not be used to quantify the differences in trends between the datasets.

Figure 7: Comparison between Trends in Standardized Life Satisfaction in the German SOEP and the Eurobarometer Survey: East Germany



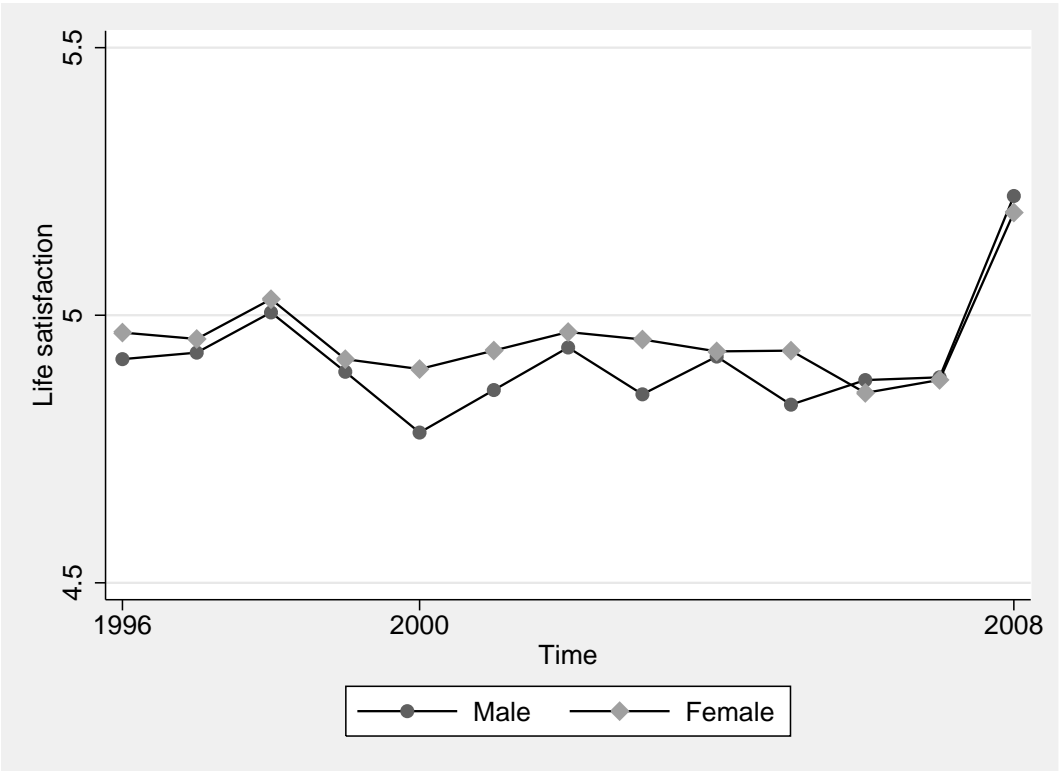
Source: German Socioeconomic Panel and Eurobarometer Survey. For both data sources, life satisfaction has been standardized with mean 0 and standard error 1. Hence, the graphs are only meant for illustrative purposes, and they should not be used to quantify the differences in trends between the datasets.

Figure 8: Trends of GHQ Scores over Time: Britain



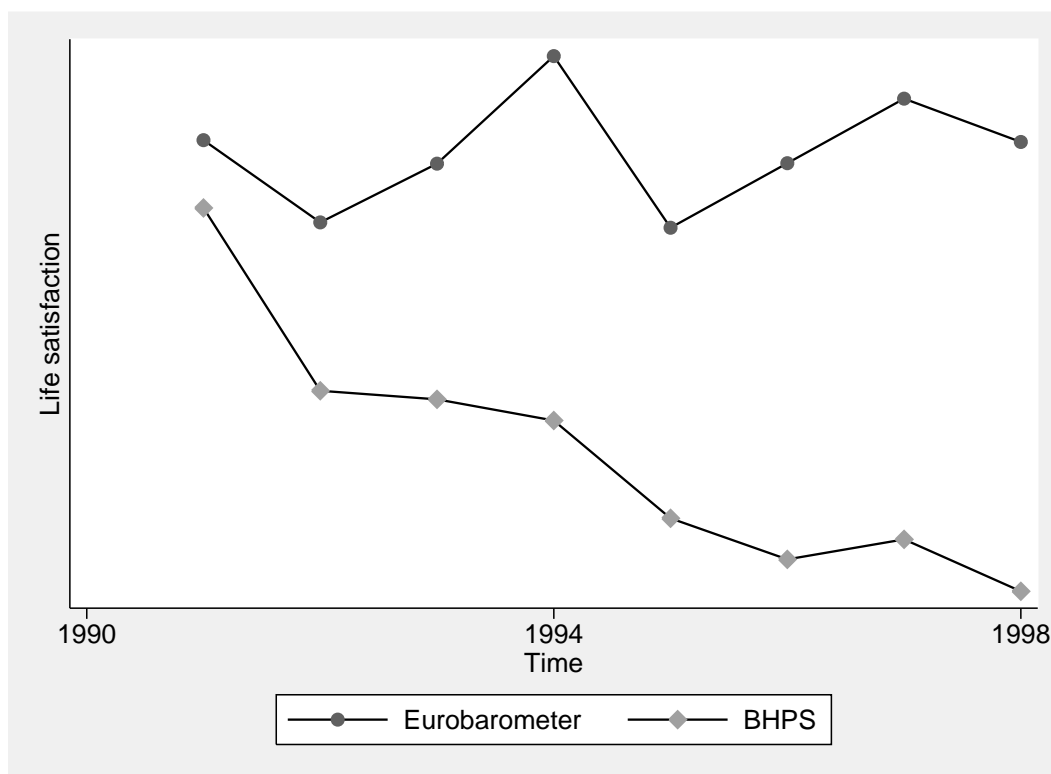
Source: British Household Panel Survey.

Figure 9: Trends of Life Satisfaction Scores over Time: Britain



Source: British Household Panel Survey.

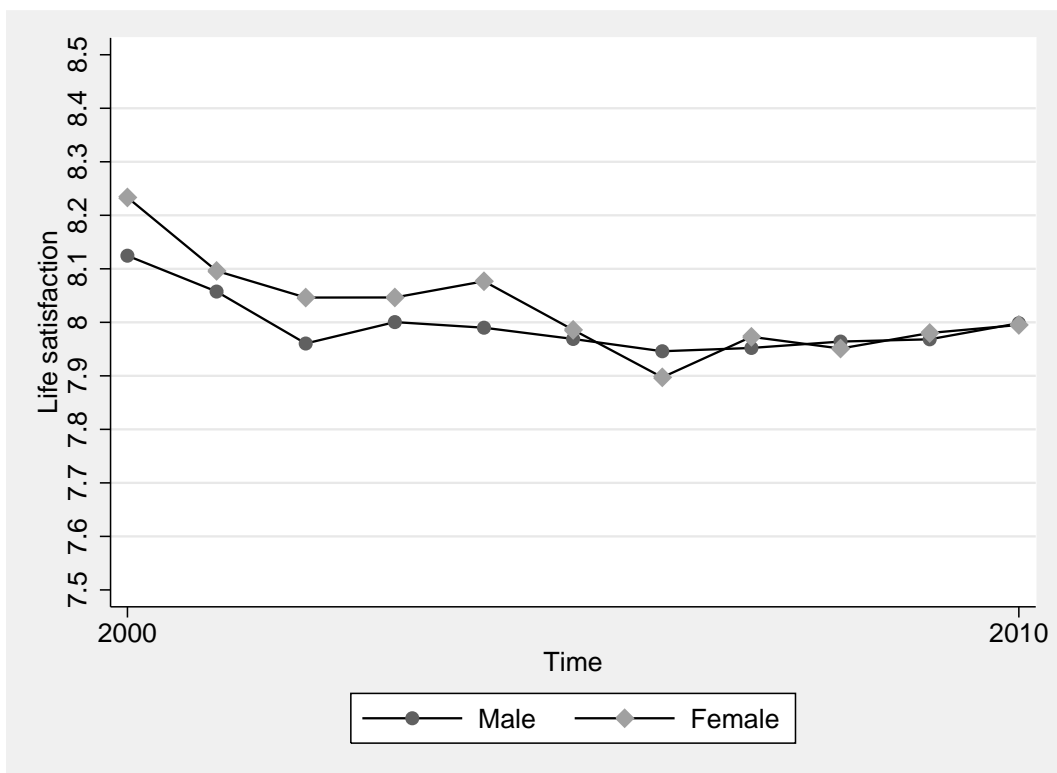
Figure 10: Comparison between Trends in Standardized GHQ Scores in the BHPS and Standardized Life Satisfaction Scores in the Eurobarometer Survey



Source: British Household Panel Survey and Eurobarometer Survey.

For both data sources, life satisfaction has been standardized with mean 0 and standard error 1. Hence, the graphs are only meant for illustrative purposes, and they should not be used to quantify the differences in trends between the datasets.

Figure 11: Trends of Life Satisfaction Scores over Time: Switzerland



Source: Swiss Household Panel.